Submitted to *Transportation Science*

manuscript TS-2020-0165.R2

Performance Evaluation of Stochastic Systems with Dedicated Delivery Bays and General On-street Parking

Abhishek

Department of Logistics, Kuehne Logistics University, Hamburg, Germany, abhishek.abhishek@the-klu.org.

Benjamin Legros

EM Normandie, Laboratoire Métis, 64 Rue du Ranelagh, 75016 Paris, France, blegros@em-normandie.fr.

Jan C. Fransoo

Tilburg School of Economics and Managemnt, Tilburg University, Netherlands, jan.fransoo@tilburguniversity.edu.

As freight deliveries in cities increase due to retail fragmentation and e-commerce, parking is becoming a more and more relevant part of transportation. In fact, many freight vehicles in cities spend more time parked than they are moving. Moreover, part of the public parking space is shared with passenger vehicles, especially cars. Both arrival processes and parking and delivery processes are stochastic in nature. In order to develop a framework for analysis, we propose a queueing model for an urban parking system consisting of delivery bays and general on-street parking spaces. Freight vehicles may park both in the dedicated bays and in general on-street parking, while passenger vehicles only make use of general on-street parking. Our model allows us to create parsimonious insights into the behavior of a delivery bay parking stretch as part of a limited length of curbside. We are able to find explicit expressions for the relevant performance measures, and formally prove a number of monotonicity results. We further conduct a series of numerical experiments to show more intricate properties that cannot be shown analytically. The model helps us shed light onto the effects of allocating scarce urban curb space to dedicated unloading bays at the expense of general on-street parking. In particular, we show that allocating more space to dedicated delivery bays can also make passenger cars better off.

Key words: Urban logistics, bay parking, street parking, parking system, loss system

1. Introduction

With the growth of cities and their population density, freight deliveries to cities are growing in volume and number while available space is more and more limited. Research on urban logistics has been growing in attention in the past decades. Recent reviews provide overviews and outlooks for challenging transportation problems in urban logistics (Malik et al. 2017, Savelsbergh and Van Woensel 2016, Taniguchi and Thompson 2018, Crainic, Ricciardi, and Storchi 2009). While most of this research addresses challenging problems to improve the actual movement of vehicles, recent empirical studies (e.g., Vieira and Fransoo 2015, Goodchild and Ivanov 2017) show that urban delivery vehicles are parked for unloading during a considerable part of the time, in some cases more than they are riding. This is particularly the case in high-density urban environments where the retail landscape (for B2B deliveries) is very dense or the consumer density (for B2C deliveries) is very high. For instance, in a field study on deliveries to nanostores (traditional convenience stores) in Latin America, Fransoo, Cedillo-Campos, and Gamez-Perez (2020) show that delivery vehicles were parked about 80% of the total time spent on the route; hence, only about 20% is spent with the vehicle in motion. Similarly high numbers have been shown in other studies in North America (Jaller, Holguín-Veras, and Hodge 2013, Goodchild and Ivanov 2017). In parcel delivery (mostly due to online commerce), vehicles in Amsterdam have been reported to be parked for about 70% of total time spent on their route. Since it is very difficult to find parking space, vehicles are searching for parking space extensively (Dowling, Ratliff, and Zhang 2019, Shoup 2006, Cassady and Kobza 1998, Dalla Chiara and Goodchild 2020) and may be forced to park illegally causing difficulties to other traffic (Gao and Ozbay 2016, Kladeftiras and Antoniou 2013). In a field experiment, Fransoo, Cedillo-Campos, and Gamez-Perez (2020) show that by providing more delivery bay space, efficiency improvements in deliveries of up to 40% could be reached. Since this efficiency gain could be used to include more cargo in the vehicle (urban delivery vehicles are often not filled to capacity), this could potentially reduce the number of vehicles needed and increase the number of deliveries per vehicle.

However, to create more dedicated space for delivery by dedicating curbside space to delivery bays, these dedicated bays take away space for general on-street parking (Nourinejad et al. 2014). Such measures generally are not very popular with car drivers and others (such as public transport and taxis) that make use of scarce curbside resources. The alternative is also not very attractive: if freight vehicles cannot find an available delivery bay, they either occupy the general-purpose parking space or park illegally. This behavior may be at the detriment to overall space utilization, as in such cases the handling time may be longer (for instance due to the fact that the space at the general on-street parking might be too small for (un)loading, or other vehicles interfere in the delivery operation). Obviously, there may also be an increased safety risk for drivers and helpers taking care of the freight delivery.

Such curbside parking stretches that have both delivery bays and general-purpose street parking spaces with stochastic arrival and service times are complex systems to model. Earlier studies (Caliskan et al. 2007, Dalla Chiara and Cheah 2017, Dowling, Ratliff, and Zhang 2019, Larson and Sasanuma 2007, Wigan and Broughton 1980, Xiao, Lou, and Frisby 2018) have investigated either freight delivery parking systems or on-street passenger vehicle (car) parking systems. However, as described above, in actual curbside parking stretches delivery vehicles make use of both delivery bays and of general on-street parking spaces. Moreover, also passenger vehicles make use of the general on-street parking spaces. Hence, a complex parking system emerges where (1) limited space is split between two types of parking spaces, (2) randomly arriving delivery vehicles that find delivery bays occupied will use general on-street parking spaces to make their deliveries, and (3) passenger vehicles make use of general on-street parking spaces. This results in a complex parking system with multiple related arrival and service processes.

In this paper, we develop such a model. The model helps us shed light onto the effects of allocating scarce urban curb space to dedicated unloading bays at the expense of general on-street parking. We consider an abstract section of a street with a given number of parking spaces, of which we allocate a certain share to delivery bays ('Bay parking'), with the remainder available for general on-street parking ('Street parking'). Using a queueing approach, we model both the bay parking stretch and the street parking stretch as a set of parallel servers with stochastic service times, and model the arrival process of new vehicles as a stochastic process. In our model, delivery vehicles have a preference for bay parking; if bay parking is not available, they will make use of street parking spaces. If these are also not available, they will leave our system. The latter in practice most likely implies that the delivery vehicle will park illegally (such as double parking or parking on the sidewalk); hence the share of delivery vehicles being blocked in our system provides an estimate of the share of delivery vehicles parking illegally. Our system also faces a stochastic arrival process of passenger vehicles that intend to park at street parking spaces, and leave our system if no place is available. We believe we are the first to model such a curbside parking system, which is very common in many cities. Our modeling approach can be the basis for much-needed analysis of the usage of scarce public space for delivery in dense cities.

From a queueing perspective, the model studied in this paper is the so-called N-design. The N-design is a canonical model that is complex to analyze since two teams of servers are considered and the overflow depends on the congestion in one of the queues. This creates a correlation between the congestion in the two parking stretches and hence a two-dimensional problem. Nevertheless, by adjusting results from the Erlang loss system (Ross 2014) to our model, we are able to derive numerically the relevant performance measures in a simple manner. Moreover, we formally prove some monotonicity results of the performance measures in the number of dedicated bay parking space. To this end, as most of the performance measures cannot be obtained in closed-form, we develop a two-dimensional iterative Markov decision process approach. Assuming that the total number of parking space is fixed, this approach allows us to prove that the bay blocking probability, the freight blocking probability, and the total utilization rate are decreasing in the number of dedicated bay parking space. This result provides an additional step in the analysis of the N-design which can contribute to the understanding of other systems modeled by this queueing architecture. These monotonicity results support our numerical observations and may be employed for deciding on the number of bay parking space.

Our model allows us to create parsimonious insights into the behavior of a delivery bay parking stretch as part of a limited stretch of curbside. We further conduct a series of numerical experiments to show more intricate properties that cannot be shown analytically. In general, and in line with common intuition, freight vehicles are better off when more curbside is reserved for dedicated delivery bays. This reduces the number of freight vehicles that are blocked to the parking system, implying fewer vehicles to park illegally. Interestingly, in freight intensive curbsides such as downtown business districts, an increase in the number of delivery bays does not only reduce the blocking probability of freight vehicles, but may also reduce the blocking probability of passenger vehicles if the freight unloading time is shorter at a bay parking stretch than at a street parking stretch. Further, in such freight intensive areas, the system blocking probability is non-monotonous in the number of delivery bays due to the complex interaction of delivery vehicles moving from blocked delivery bays to available on-street parking spaces. Hence, a system may be performing better overall, if the number of delivery bays is increased. This may seem counterintuitive, since common intuition may be that fewer allocation restrictions would improve overall performance. In such freight-intensive systems, it is hence critical for urban planners to explicitly model the interaction between the usage of dedicated delivery bays and general on-street parking by delivery vehicles. Finally, we show that in passenger intensive areas, the interaction between the bay parking stretch and the street parking stretch is much more limited, since the street parking spaces are usually occupied by cars. Due to this limited interaction, metrics like system utilization are much less sensitive to delivery vehicles making use of the general street parking. In such systems, urban planners can analyze the consequences of both spaces more or less independently.

We believe our model is the first to study the role of delivery bays in an urban logistics setting in a stochastic manner. The framework we provide can serve as a basis for further work in this area, as it can be extended to a queueing network to represent more extensive relations between multiple separate bay parking stretches and street parking stretches in an urban setting.

The remainder of our paper is organized as follows. Section 2 reviews the literature. Section 3 defines the queueing model under consideration. Section 4 analyzes our model and gives the performance measures and their monotonocity properties in the number of parking spaces. Section 5 presents an approximate analysis for the parking system allowing us to study the effect of different parking times between delivery vehicles and passenger vehicles, and provides an approximation to incorporate time-dependent vehicle arrivals in our modeling framework. Section 6 presents numerical examples to demonstrate the impact of explicitly linking the bay parking stretch with the street parking stretch on the blocking probability of vehicles and the system utilization. We conclude in Section 7.

2. Literature Review

In the existing literature, various queueing models have been studied for parking systems, among which the $M/M/c/c$ queue is the most widely used model (Hauer and Templeton 1972, Caliskan et al. 2007, Dalla Chiara and Cheah 2017, Dowling, Ratliff, and Zhang 2019). Wigan and Broughton (1980) use a discrete-time semi-Markov chain to investigate the occupancy of a parking area over time by allowing variations in the arrival of vehicles and their lengths of stay. Based on a birth-death process, a queueing model is studied to capture the behavior of cruising drivers searching for inexpensive on-street parking, where economic effects of congestion pricing are analyzed (Larson and Sasanuma 2007). Caliskan et al. (2007) develop a model using a continuous-time homogeneous Markov process to predict parking lot occupancy at the time of arrival in a vehicular ad-hoc network (VANET). Based on the parking information exchanged among vehicles through the VANET, their model enables each vehicle to make a parking decision on the available parking lots. Dalla Chiara and Cheah (2017) investigate loading bays at urban malls as queueing systems. In their study, they collect data through various means to first estimate the arrival rates and the parking duration of freight vehicles unloading goods at two large urban retail malls in Singapore, and then analyze the congestion of the loading bays and its effect on the drivers' choice of parking. Using an $M/M//c/c$ queueing model, Xiao, Lou, and Frisby (2018) propose a model-based predictive framework for parking occupancy, which they validate using both real and simulated data. In addition, they show numerically that their proposed model-based predictive method performs better than pure machine learning parking occupancy prediction methods. In a recent study, Dowling, Ratliff, and Zhang (2019) examine curbside parking as a network of finite capacity queues, where performance metrics such as the rate of rejection of vehicles searching for parking and the parking occupancy are estimated using a simulation. In our paper, we consider both freight vehicles conducting delivery activities *and* passenger vehicles seeking parking space. They make use of the same common parking space, where a share of this parking space has been reserved for loading and unloading *only*. We believe such a system – despite being very common in the reality of cities – has not been studied in the literature. Especially the interaction between freight vehicles and passenger vehicles makes it very different from any of the prior studies.

From a more general perspective beyond parking, the queueing model studied in this paper is designed to provide an *overflow mechanism*. In the queueing literature, overflow policies have been widely studied. The most simple overflow strategy is to reject or outsource (in the context of call centers) some customers at arrival from a single queue (Ku and Jordan 2003, Maglaras and Van Mieghem 2005, Ward and Kumar 2008, Xu 2015, Niyirora and Zhuang 2017). These papers study the relation between the wait and the rejection flow. The system's design parameters are the staffing level and the rejection threshold. However, in these prior studies the service quality offered to rejected customers is neglected. Therefore, as in our paper, other studies propose to build queueing models which provide service to customers in overflow. For this purpose, more complex architectures have been investigated. For instance, in the context of outsourcing, Gans and Zhou (2007) study a call center with high and low value calls and evaluate routing schemes for outsourcing part of the low value calls, investing different priority queues. Gurvich and Perry (2012) consider a service network operated under a threshold-type overflow mechanism. If the waiting room is full, the call is overflowed to an outsourcer. They show that the larger the system becomes, the more negligible the dependency between each in-house station and overflow station. The same phenomenon is also observed in our queueing model. Specifically, the architecture in our paper is the so-called N-design. The N-design is more complex than the V-design as two teams of servers are considered and the overflow depends on the congestion in one of the queues (Bassamboo, Harrison, and Zeevi 2006). Although the performance in the N-design model can only be obtained numerically, using a Markov decision process approach Koole, Nielsen, and Nielsen (2015) tackle the problem of optimizing customers' overflow based on the system's wait. In contrast with their study, the theoretical novelty in our paper is that we investigate the monotonicity properties in the staffing levels for the N-design. Our study hence is the first to explicitly study the performance for freight delivery of a parking system that is very common in many cities: a collection of dedicated delivery bays and general on-street parking spaces. In order to model this, we develop a novel queueing model that bares some similarities to earlier models in the call-center performance analysis literature. We theoretically show monotonicity properties of our parking system, and numerically analyze its performance to help urban planners better understand the trade-offs when allocating scarce public parking space.

3. Model

We consider a parking system of c parking spaces with two types of vehicle; freight delivery vehicles and passenger vehicles. We refer to the former ones as class-1 vehicles while the latter are class-2 vehicles. The arrival process of class-i vehicles is Poisson with parameter λ_i for $i = 1, 2$. The parking system consists of c_b delivery bays (bay parking) and c_s general on-street parking spaces (street parking) with $c_s + c_b = c$. Freight vehicles (i.e. class-1 vehicles), arriving for unloading, are routed in priority to the bay parking stretch. If they find vacant parking spaces, they enter the bay parking stretch, complete their deliveries, and then leave the system. If all bay parking spaces are occupied, they move to the street parking stretch in order to find vacant parking spaces so that they can make their deliveries at the street parking stretch. If also no vacant parking spaces at the street parking stretch are available (i.e., all parking spaces are occupied), they leave the parking system. Passenger vehicles (i.e., class-2 vehicles) are only allowed to park at the street parking stretch.

We assume that parking times at the bay and street parking are exponentially distributed with parking rates μ_b and μ_s, respectively. In this way, parking times have the same distribution for class-1 and class-2 vehicles at the street parking and class-1 vehicles have different parking times whether they park at the bay or at the street parking. While the former assumption is made for tractability, the latter one allows parking times (service times) of delivery vehicles at the street parking stretch to be longer than that at the bay parking stretch. This is due to the fact that the space at the street parking stretch might be too small for (un)loading, or other vehicles interfere in the delivery operation. Nevertheless, parking times at the street parking might be different for class-1 and class-2 vehicles. The parking system is depicted in Figure 1. This initial model is analyzed is Section 4. We

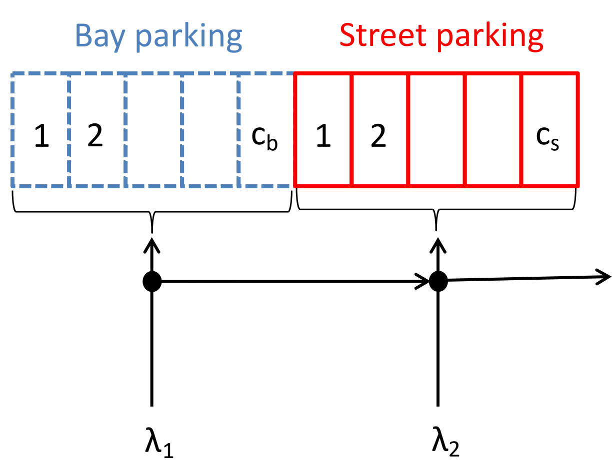


Figure 1 The parking viewed as a queueing system

extend this model in Section 5 and develop approximations to account for different parking times for class-1 and class-2 vehicles at the street parking, termed μ_s^1 and μ_s^2, and time-dependent sinusoidal arrival rates, termed $\lambda_i(t)$ for $t \geq 0$.

External validation. We validate our key assumptions on the distributions through a statistical analysis. Specifically, we conducted an empirical data analysis supporting our assumption that the arrival processes of both freight and passenger vehicles follow Poisson processes with different rate parameters, and their parking times are exponentially distributed. We use the 2014 data from the city of Melbourne, Australia (Melbourne City Council 2014) that has been collected using sensors on parking bays in the downtown area. In addition, the dataset includes specific information whether a parking bay is a delivery bay (for freight) or a general parking area (for general on-street parking). We use all transactional data for one particular parking system consisting of a stretch of 6 delivery bays and 6 general on-street parking places, and analyze these transactional data in one-hour clusters for freight parking and two-hours clusters for passenger parking. For each of these hours, we test our distribution assumptions using the appropriate Kolmogorov-Smirnov (KS) goodness-of-fit test (Kolmogorov 1933, Smirnov 1948). For all but one of the delivery vehicle arrival time clusters and also for all but one of the car arrival time clusters, we do not exceed the KS statistic and hence cannot reject the hypothesis that the distribution of the parking times is exponential. We use the same data but clustered by day to test the hypothesis that the arrivals are Poisson distributed. With 174 days of observations with freight vehicles, in 85% of the days we cannot reject the hypothesis that the arrival process is Poisson. Similar, for passenger vehicles this is true for 89%. It hence seems quite reasonable to make the assumptions as stated above.

Performance measures. To analyse the parking system, we define some key stationary performance measures which allows us to understand the parking system behavior and the service quality offered to users. To measure service quality we consider the blocking probability, which is the probability of not finding an available parking space at arrival. Class-1 vehicles have access to both parking lots. We thus distinguish between the blocking probability of class-1 vehicles given an arrival at the bay and at the street parking, termed π_1^b and π_1^s, respectively. The total blocking probability for class-i vehicles is termed π_i for $i = 1, 2$ and the total blocking probability of an arbitrary vehicle is termed π.

The system behavior is characterized by its utilization, defined as the expected proportion of occupied parking spaces. We consider the bay and street parking expected utilization, termed U_b and U_s, respectively, and the global parking system utilization, termed U. To define different workload levels, we use the notion of offered load per server. It is computed as the ratio between the arrival rate in one system divided by the product of the number of parking spaces and the parking rate. Thus, an offered load per server of 1 or more corresponds to a system operating at its maximum processing capacity where the mean rate of vehicles arrivals is equal to or exceeds the processing capacity of the parking. As for the expected utilization, we consider the offered load per server at the bay parking, termed ρ_b, at the street parking, termed ρ_s.

Table of Notations. We end this section with a table of Notations (Table 1) used throughout the paper.

Table 1 Table of notations

System parameters	
λ_i	Arrival rate of class-i vehicles for $i=1,2$
μ_b	Parking rate at the bay parking (i.e. $1/\mu_b$ is the expected parking time at the bay parking)
μ_s	Parking rate at the street parking (i.e. $1/\mu_s$ is the expected parking time at the street parking)
μ_s^i	Parking rate of class-i vehicles at the street parking for $i=1,2$ when class-1 and class-2 vehicles have different parking times
c_b	Capacity of the bay parking
c_s	Capacity of the street parking
c	Total system capacity (i.e., $c=c_b+c_s$)
Performance measures	
U_b, U_s, U	Bay utilization, street utilization, system utilization
ρ_b, ρ_s	Offered load per server at the bay parking and at the street parking, respectively
π_1^b	Blocking probability of a class-1 vehicle given an arrival at the bay parking
π_1^s	Blocking probability of a class-1 vehicle given an arrival at the street parking
π_1	Blocking probability of a class-1 vehicle at an arrival instant in both parking lots
π_2	Blocking probability of a class-2 vehicle
π	Blocking probability of an arbitrary vehicle

4. Analysis

We now provide an analysis of the queueing system introduced above. In Section 4.1, we derive the main performance measures. In Section 4.2, we provide key monotonicity results to explain the behavior of our queueing system related to the allocation of parking spaces between bay parking and street parking. Theoretically, the monotonicity provides important insights into the fundamental characteristics of the system. From a more pragmatic perspective, the monotonicity results allow for the creation of simple algorithms that can be deployed for decision making in urban planning.

4.1. Performance measures

We obtain the performance measures from the analysis of the $M/M/c_b/c_b$ and $GI+M/M/c_s/c_s$ queues which correspond to the bay and street parking, respectively. The performance measures of interest are presented in (1)-(8).

Bay parking. The bay parking stretch can be viewed as an $M/M/c_b/c_b$ queue (Ross 2014), where the first M in the Kendall notation denotes the Markovian interarrival times with rate λ_1, the second M refers to the exponential parking times with rate μ_b, the first c_b is the total number of bay parking spaces, and the last c_b refers to the capacity of the bay parking. This queueing model is also known as the Erlang loss system. We recall the performance measures of interest for this model (e.g., see p.564 in Ross (2014)):

$$\pi_1^b = \left[\sum_{k=0}^{c_b} \frac{\left(\frac{\lambda_1}{\mu_b}\right)^{k-c_b} c_b!}{k!} \right]^{-1}, \quad U_b = \frac{\lambda_1}{c_b\mu_b}(1-\pi_1^b), \text{ and } \rho_b = \frac{\lambda_1}{c_b\mu_b}. \tag{1}$$

Street parking. One useful result for the analysis is that the output flow of lost class-1 vehicles from an $M/M/c_b/c_b$ queue forms a renewal process. This result is stated in the first paragraph of Section 3 in Takács (1959). An overview of the results related to the output flows in standard queueing systems can be found in Daley (1976). Therefore, both types of vehicles arrive at the street parking stretch according to mixed renewal and Poisson processes (denoted by GI+M). Thus, the street parking stretch is considered as a $GI+M/M/c_s/c_s$ queue. As a consequence, the total arrival rate of vehicles (including both freight and passenger) to the street parking stretch, is equal to $\lambda_1\pi_1^b+\lambda_2$. Consequently, the offered load per server at the street parking stretch is given by

$$\rho_s = \frac{\lambda_1\pi_1^b+\lambda_2}{\mu_s c_s}. \tag{2}$$

The other performance measures cannot be obtained in closed-form. Using the $GI+M/M/c_s/c_s$ queueing model analyzed by Kuczura (1973), we numerically derive the steady-state probability of occupying j parking spaces at an arrival epoch of a class-1 vehicle, denoted by p_j^1 for $j=0,1,\cdots,c_s$, by solving the following linear system of equations:

$$p_j^1 = \sum_{i=0}^{c_s} p_i^1 r_{i,j}, \quad j=0,1,\cdots,c_s,$$

where

$$r_{i,j} = a_j + \sum_{k=1}^{c_s} \beta_k(i+1,j)\tilde{\alpha}(-\gamma_k\mu_s), \text{ for } i=0,1,\cdots,c_s-1,$$
$$r_{c_s,j} = a_j + \sum_{k=1}^{c_s} \beta_k(c_s,j)\tilde{\alpha}(-\gamma_k\mu_s),$$

with

$$\tilde{\alpha}(s) = \frac{\sum_{j=0}^{c_b}\binom{c_b}{j}\prod_{i=0}^{j-1}\frac{s+i\mu_b}{\lambda_1}}{\sum_{j=0}^{c_b+1}\binom{c_b+1}{j}\prod_{i=0}^{j-1}\frac{s+i\mu_b}{\lambda_1}}, \quad a_j = \left[\sum_{k=0}^{c_s}\frac{\left(\frac{\lambda_2}{\mu_s}\right)^{k-j} j!}{k!}\right]^{-1}, \quad \beta_k(i,j) = \frac{c_s!\left(\frac{\lambda_2}{\mu_s}\right)^{c_s-i} D_i(\gamma_k)D_j(\gamma_k)}{j!\gamma_k D_{c_s}(\gamma_k)D'_{c_s}(\gamma_k+1)},$$

and where the γ_k's, $k=1,2,\cdots,c_s$, are the k roots of the polynomial $D_{c_s}(\xi+1)$, with $D_i(\xi)=\sum_{k=0}^{i}\binom{i}{k}\left(\frac{\lambda_2}{\mu_s}\right)^{i-k}\xi(\xi+1)\cdots(\xi+k-1)$ for $i>0$ and $D_0(\xi)=1$.

The blocking probability of a class-1 vehicle in the street parking stretch is the probability of finding the fully occupied street parking stretch at its arrival time. Therefore, we have

$$\pi_1^s = p_{c_s}^1. \tag{3}$$

There remains to find the blocking probability of passenger vehicles in the street parking stretch. The steady-state probability of occupying j servers at an arrival epoch of a class-2 vehicle, denoted by p_j^2, is given by

$$p_j^2 = \frac{\left(\frac{\lambda_2}{\mu_s}\right)^j p_0^2}{j!} + \frac{\lambda_1 \pi_1^b}{\mu_s}\sum_{i=0}^{j-1}\frac{i!\left(\frac{\lambda_2}{\mu_s}\right)^{j-i-1} p_i^1}{j!}, \text{ for } j=1,2,\cdots,c_s,$$

where p_0^2 is determined by the normalization condition, i.e.,

$$p_0^2 = \frac{1-\frac{\lambda_1\pi_1^b}{\lambda_2}\sum_{j=1}^{c_s}\frac{\left(\frac{\lambda_2}{\mu_s}\right)^j}{j!}\sum_{i=0}^{j-1}\frac{i!p_i^1}{\left(\frac{\lambda_2}{\mu_s}\right)^i}}{\sum_{j=0}^{c_s}\frac{\left(\frac{\lambda_2}{\mu_s}\right)^j}{j!}}.$$

Thus, the blocking probability of arriving passenger vehicles (class-2 blocking probability) in the street parking stretch is given by

$$\pi_2 = p_{c_s}^2. \tag{4}$$

The total blocking probability of an arbitrary arriving vehicle in the street parking stretch is the weighted average of the blocking probabilities of class-1 and class-2 vehicles in the street parking stretch: $\frac{\lambda_2}{\lambda_2+\lambda_1\pi_1^b}\pi_2+\frac{\lambda_1\pi_1^b}{\lambda_2+\lambda_1\pi_1^b}\pi_1^s$. Therefore, the total effective arrival rate of class-1 and class-2 vehicles to the street parking stretch equals $(\lambda_2+\lambda_1\pi_1^b)\left(1-\frac{\lambda_2}{\lambda_2+\lambda_1\pi_1^b}\pi_2-\frac{\lambda_1\pi_1^b}{\lambda_2+\lambda_1\pi_1^b}\pi_1^s\right)$. Using Little's law (Little 1961), the expected number of occupied parking spaces at the street parking stretch, is equal to the effective arrival rate multiplied by the mean time that a vehicle spends in the system. By dividing this quantity by c_s, we deduce the street utilization:

$$U_s = \frac{\lambda_2+\lambda_1\pi_1^b}{\mu_s c_s}\left(1-\frac{\lambda_2}{\lambda_2+\lambda_1\pi_1^b}\pi_2-\frac{\lambda_1\pi_1^b}{\lambda_2+\lambda_1\pi_1^b}\pi_1^s\right). \tag{5}$$

Parking system. As the class-1 vehicles arriving to the street parking stretch are those which are blocked to the bay parking stretch, the total blocking probability of class-1 vehicles in the parking system is the product of the blocking probabilities of freight vehicles in the bay parking stretch and in the street parking stretch:

$$\pi_1 = \pi_1^s \pi_1^b. \tag{6}$$

The total blocking probability of an arbitrary arriving vehicle in the parking system (system blocking probability) is the weighted average of the blocking probabilities of class-1 and class-2 vehicles:

$$\pi = \frac{\lambda_1}{\lambda_1+\lambda_2}\pi_1 + \frac{\lambda_2}{\lambda_1+\lambda_2}\pi_2. \tag{7}$$

Similarly, the utilization of the parking system (system utilization) at an arbitrary time is the weighted average of the bay utilization and the street utilization at an arbitrary time:

$$U = \frac{c_b}{c}U_b + \frac{c_s}{c}U_s. \tag{8}$$

4.2. Monotonicity properties of the performance measures

We now investigate the monotonicity properties in c_b of the performance measures given in (1)-(8). This analysis supports the numerical investigations of Section 6. As some of the performance measures cannot be obtained in closed-form, we pursue an iterative Markov decision process approach where monotonicity results can be proven without using explicit formulas. We hold the total number of parking spaces c constant. Given that $c_b + c_s = c$, having one extra space at the bay parking stretch means removing one parking space from the street parking stretch.

The Markov chain. Our system can be represented by a finite state Markov chain. A state of the system is denoted by (x, y), where x is the number of vehicles at the bay parking stretch for $0 \le x \le c_b$ and y is the number of vehicles at the street parking stretch for $0 \le y \le c - c_b$. We do not need to distinguish between class-1 and class-2 vehicles at the street parking stretch since they both have the same parking time distribution. The four possible transitions in the Markov process are as follows:

- A class-1 vehicle arrival with rate λ_1 from state (x, y) with $y < c - c_b$, which changes the state either to $(x+1, y)$ if $x < c_b$, that is the number of vehicles at the bay parking stretch is increased by one, or to $(c_b, y+1)$ if $x = c_b$, that is the number of vehicles at the street parking stretch is increased by one.

- A class-2 vehicle arrival at the street parking stretch with rate λ_2 from state (x,y) with $y < c - c_b$, which changes the state to $(x, y+1)$, that is the number of vehicles at the street parking stretch is increased by one.
- A parking lot exit with rate $x\mu_b$ from state (x,y) with $x > 0$, which changes the state to $(x-1, y)$, that is the number of vehicles at the bay parking is decreased by one.
- A parking lot exit with rate $y\mu_s$ from state (x,y) with $y > 0$, which changes the state to $(x, y-1)$, that is the number of vehicles at the street parking is decreased by one.

The cost function. We introduce a cost function $f(c_b, x, y)$ for $0 \le x \le c_b$ and $0 \le y \le c - c_b$, which enables us to compute any desired performance measure, except π_1^s. Given that the arrival process of class-1 vehicle at the street parking is not Poisson, we cannot capture π_1^s via a Markov decision process formulation. In Table 2, we indicate how the cost function $f(c_b, x, y)$ should be defined to capture a desired metric. We denote by $\mathbb{1}_A$ the indicator function of a given subset A.

Table 2 Definitions of the cost functions $f(c_b, x, y)$

Performance measures	$f(c_b, x, y)$
π_1^b	$\mathbb{1}_{x=c_b}$
π_1	$\mathbb{1}_{x=c_b, y=c-c_b}$
π_2	$\mathbb{1}_{y=c-c_b}$
π	$\frac{\lambda_1 \mathbb{1}_{x=c_b, y=c-c_b} + \lambda_2 \mathbb{1}_{y=c-c_b}}{\lambda_1+\lambda_2}$
U_b	$\frac{x}{c_b}$
U_s	$\frac{y}{c-c_b}$
U	$\frac{x+y}{c}$
ρ_b	$\frac{\lambda_1}{\mu_b c_b}$
ρ_s	$\frac{\lambda_1 \mathbb{1}_{x=c_b} + \lambda_2}{\mu_s (c-c_b)}$

Computation of the value function. Our system is uniformizable as the maximal event rate, $\lambda_1 + \lambda_2 + c_b\mu_b + (c-c_b)\mu_s$, is bounded. We assume, without loss of generality that $\lambda_1 + \lambda_2 + c_b\mu_b + (c-c_b)\mu_s = 1$, such that the transition rates in the continuous time Markov chain are viewed as transition probabilities in the discrete time one. We are now in position to introduce the system value function $V_k(c_b, x, y)$ over k steps for $0 \le x \le c_b$, $0 \le y \le c - c_b$, and $k \ge 0$. For $k \ge 0$, $0 \le x \le c_b$, and $0 \le y \le c - c_b$, we choose $V_0(c_b, x, y) = 0$, and we express V_{k+1} as a function of V_k as follows:

$$\begin{aligned} V_{k+1}(c_b, x, y) = {} & f(c_b, x, y) \\ & + \lambda_1 \left(\mathbb{1}_{x<c_b} V_k(c_b, x+1, y) + \mathbb{1}_{x=c_b, y<c-c_b} V_k(c_b, x, y+1) + \mathbb{1}_{x=c_b, y=c-c_b} V_k(c_b, x, y) \right) \\ & + \lambda_2 \left(\mathbb{1}_{y<c-c_b} V_k(c_b, x, y+1) + \mathbb{1}_{y=c-c_b} V_k(c_b, x, y) \right) \\ & + x\mu_b V_k(c_b, x-1, y) + y\mu_s V_k(c_b, x, y-1) + (1 - \lambda_1 - \lambda_2 - x\mu_b - y\mu_s) V_k(c_b, x, y). \end{aligned} \tag{9}$$

As k tends to infinity, the difference $V_{k+1} - V_k$ converges to the long-run average cost (Puterman 2014), that is the desired performance measure. The convergence is due to the aperiodic irreducible finite-state Markov chain considered here. The aperiodicity is due to the fictitious transitions from a state to itself. Therefore, by studying the properties of V_k in c_b, we can deduce those of the performance measures. This approach has been successfully employed to show monotonicity results for some queueing model where either the performance measures cannot be found explicitly (Bhulai, Brooms, and Spieksma 2014) or they lead to formulas which are difficult to manipulate (Legros and Jouini 2019, Legros 2021).

Monotonicity results. In Corollary 1, we provide the first order monotonicity properties of the performance measures. To this end, in Theorem 1, we prove by induction on k that V_k belongs to some classes of functions, $\mathcal{C}$ and $\mathcal{C}'$. We define the class of functions $\mathcal{C}$ as follows: $h \in \mathcal{C}$ if for $0 \leq x \leq c_b \leq c$, and $0 \leq y \leq c-(c_b+1)$, we have

$$h(c_b, x, y) \geq h(c_b+1, x, y), \tag{10}$$

$$h(c_b, x, y+1) \geq h(c_b, x, y), \text{ and} \tag{11}$$

$$h(c_b, x, y) \geq h(c_b+1, x+1, y). \tag{12}$$

We also define the class of functions $\mathcal{C}'$ which differs from $\mathcal{C}$ where (12) is replaced by (13), defined as

$$h(c_b, x, y+1) \geq h(c_b+1, x+1, y), \tag{13}$$

for $0 \leq x \leq c_b \leq c$, and $0 \leq y \leq c-(c_b+1)$. The class of function $\mathcal{C}'$ is a subset of $\mathcal{C}$. This can be seen by summing up (11) and (12), leading to (13). By summing up (10) and (11), we generate (14):

$$h(c_b, x, y+1) \geq h(c_b+1, x, y), \tag{14}$$

for $0 \leq x \leq c_b \leq c$, and $0 \leq y \leq c-(c_b+1)$. This relation is satisfied both in $\mathcal{C}'$ and $\mathcal{C}$. Relation (10) for V_k proves that a performance measure is decreasing in c_b. Relations (11) and (12) are needed to prove that (10) holds for V_k. However, for most definitions of the cost function $f(c_b, x, y)$, (12) is not satisfied. That is why we alternatively introduce (13) which is met for more cost functions. However, to show (13) without (12) we need to have $\mu_b \geq \mu_s$.

Theorem 1 *The following holds:*

- *If $f \in \mathcal{C}$, then $V_k \in \mathcal{C}$, for $k \geq 0$.*

- *If $f \in \mathcal{C}'$ and $\mu_b \geq \mu_s$, then $V_k \in \mathcal{C}'$, for $k \geq 0$.*

In Table 3, we indicate whether f satisfies Relations (10)-(13) for the different performance measures under consideration. In case either f belongs to $\mathcal{C}'$ or $\mathcal{C}$, then the considered performance measure is decreasing in c_b. The proof of Corollary 1 follows from Table 3.

Table 3 Properties of the cost function

Cost function $f(c_b, x, y)$	Relation (10)	Relation (11)	Relation (12)	Relation (13)
$\mathbb{1}_{x=c_b}$ for π_1^b	Yes	Yes	Yes	Yes
$\mathbb{1}_{x=c_b,y=c-c_b}$ for π_1	Yes	Yes	No	Yes
$\mathbb{1}_{y=c-c_b}$ for π_2	No	Yes	No	Yes
$\frac{\lambda_1 \mathbb{1}_{x=c_b,y=c-c_b} + \lambda_2 \mathbb{1}_{y=c-c_b}}{\lambda_1+\lambda_2}$ for π	No	Yes	No	Yes
$\frac{x}{c_b}$ for U_b	Yes	Yes	No	No
$\frac{y}{c-c_b}$ for U_s	No	Yes	No	Yes
$\frac{x+y}{c}$ for U	Yes	Yes	No	Yes
$\frac{\lambda_1}{\mu_b c_b}$ for ρ_b	Yes	Yes	Yes	Yes
$\frac{\lambda_1 \mathbb{1}_{x=c_b} + \lambda_2}{\mu_s(c-c_b)}$ for ρ_s	No	Yes	No	No

Corollary 1 *The following holds:*

- *The bay blocking probability π_1^b and the offered load per server at the bay parking ρ_b are decreasing in c_b.*
- *Under the condition $\mu_b \geq \mu_s$, the freight blocking probability π_1 and the total utilization rate U are decreasing in c_b.*

The results of Corollary 1 are novel in the analysis in the $N-$design queueing model. In addition, these results may allow us to dimension the two parking systems by setting the number of parking spaces at the bay parking stretch c_b. The value of c_b is determined to ensure a certain number of dedicated parking spaces for freight vehicles only. Given that the total number of parking spaces in the parking system c is kept constant, having a large c_b worsens the performance metrics for passenger vehicles.

Example: For instance, one objective of the system manager may be to select c_b as small as possible such that the freight blocking probability, π_1, remains below a certain service level objective, $\overline{\pi_1}$ (i.e., $\pi_1 \leq \overline{\pi_1}$). From Corollary 1, π_1 is decreasing in c_b. Thus, the derivation of c_b can be obtained after a finite number of iterations from a simple algorithm stated as follows:

- Step 0: Set $c_b = c$ and compute π_1. If $\pi_1 > \overline{\pi_1}$, then the algorithm stops and the constraint $\pi_1 \leq \overline{\pi_1}$ cannot be met. Otherwise, go to Step 1.
- Step 1: Set $c_b = 0$ and compute π_1. If $\pi_1 \leq \overline{\pi_1}$, then $c_b = 0$ is optimal, otherwise go to Step 2.

- Step 2: Set $c_b := c_b + 1$ and compute π_1. If $\pi_1 \leq \overline{\pi_1}$, then c_b is optimal, otherwise go back to Step 2.

Regardless of the evidence resulting from our data analysis, it is interesting and relevant to note that our analysis above is robust to the distribution assumption of the parking times. That may be surprising, because the approximation of a general distribution by an exponential one for the classical single-server M/GI/1 queue would lead to a very poor approximation quality. The particularity of our model, however, is that waiting is not permitted. The *bay parking* behaves exactly as the Erlang loss model. In the Erlang loss model, the steady-state distribution does not depend on a general service-time distribution beyond its mean.

5. Incorporating nonidentical parking times and time-dependent arrivals

In this section, we propose two approximations to obtain the performance measures with nonidentical parking time distributions at the street parking for class-1 and class-2 vehicles in Section 5.1 and for time-dependent sinusoidal arrival rates in Section 5.2. The approximations are evaluated in Section 5.3.

5.1. Nonidentical freight and passenger vehicle parking time distributions

In the initial model analyzed in Section 3, parking times of class-1 and class-2 vehicles are exponentially distributed with the same rate μ_s. In reality, however, it is likely that the parking time distributions between freight and passenger vehicles in the street parking stretch are nonidentical. To account for this, we assume that the parking time a class-i vehicle at the street parking is exponentially distributed with rate μ_s^i for $i = 1, 2$ with $\mu_s^1 \neq \mu_s^2$. this means that the parking time at the street parking has an hyperexponential distribution with rates μ_s^1 and μ_s^2.

The idea of the approximation is to assume that the parking rate of an arbitrary vehicle at the street parking is the mean parking rate between class-1 and class-2 vehicles. Therefore, we replace the hyperexponential parking time distribution by an exponential one. In this way, the analysis of Section 4.1 can be reemployed to derive the performance measures. As the arrival rate of class-1 and class-2 vehicles at the street parking are $\lambda_1 \pi_1^b$ and λ_2, respectively, the mean parking rate μ_s is given by

$$\frac{1}{\mu_s} = \frac{\lambda_1 \pi_1^b}{\lambda_1 \pi_1^b + \lambda_2} \frac{1}{\mu_s^1} + \frac{\lambda_2}{\lambda_1 \pi_1^b + \lambda_2} \frac{1}{\mu_s^2}.$$

Note that this approximation does not impact the performance measures at the bay parking stretch.

5.2. Time-dependent arrivals

In our model in Section 3, the arrival rates at each parking stretch are assumed to be constant over time. This may not be realistic, because in some areas, there is a very pronounced variation depending on time-of-day or day-of-week. If the time-dependency is slowly varying relative to the system dynamics, then such systems have been typically analyzed using a point-wise stationary approximation, where the performance at time t is approximated by the steady-state performance of the stationary system with constant arrival rates given by the mean arrival rate on a given interval around the observation point (Green and Kolesar 1991, Jennings et al. 1996). We propose to employ this approach to approximate the long-run performance measures.

We assume that the arrival rates are sinusoidal time-dependent parameters as in Eick, Massey, and Whitt (1993), with

$$\lambda_i(t) = \overline{\lambda_i} + \overline{\lambda_i}\alpha_i \sin\left(\frac{2\tilde{\pi}t}{\Psi_i}\right),$$

for $i = 1, 2$, with $\overline{\lambda_i} > 0$, $0 < \alpha_i < 1$ and where $\tilde{\pi} \simeq 3.14159265$. The expression of $\lambda_i(t)$ is convenient for interpretation; $\overline{\lambda_i}$ is the average arrival rate, α_i is the relative amplitude, and Ψ_i is the cycle length. The mean arrival rate over an interval $[t, t+L]$, $\lambda_i^{[t,t+L]}$, is given by

$$\begin{aligned}\lambda_i^{[t,t+L]} &= \frac{1}{L}\int_{z=t}^{t+L}\left[\overline{\lambda_i} + \overline{\lambda_i}\alpha_i \sin\left(\frac{2\tilde{\pi}z}{\Psi_i}\right)\right]\,\mathrm{d}z \\ &= \overline{\lambda_i} + \frac{\overline{\lambda_i}\alpha_i}{L}\frac{\Psi_i}{\tilde{\pi}}\sin\left(\frac{\tilde{\pi}(2t+L)}{\Psi_i}\right)\sin\left(\frac{\tilde{\pi}L}{\Psi_i}\right), \text{ for } i = 1, 2.\end{aligned}$$

In our approximation, we consider intervals of equal length L, where L is chosen such that the number of intervals per cycle, $n_1 = \Psi_1/L$ and $n_2 = \Psi_2/L$, are integers. To obtain a full cycle common to class-1 and class-2 arrivals, we need to observe a number of intervals equal to the least common multiple of n_1 and n_2. Using the mean arrival rate on each interval, we determine the performance measures of each interval using the results of our original model as presented in Section 4.1. To estimate the long-run performance measures, we compute the average performance measures weighted by the mean arrival rate on each interval.

5.3. Evaluation of the approximations quality

We now evaluate the quality of the proposed approximations. In Table 4, we derive the maximal error between simulation and the approximation for each performance measure by varying c_b from 0 to c. The error is computed as the difference between the performance obtained via a simulation and the one obtained with the approximation. We present the following sets of parameters with non-identical parking times:

- Set 1: $c = 20, \lambda_1 = 0.8, \lambda_2 = 0.4, \mu_b = \mu_s^1 = 1/30, \mu_s^2 = 1/60$,
- Set 2: $c = 20, \lambda_1 = 0.8, \lambda_2 = 0.4, \mu_b = \mu_s^1 = 1/30, \mu_s^2 = 1/120$,
- Set 3: $c = 20, \lambda_1 = 0.2, \lambda_2 = 0.3, \mu_b = \mu_s^1 = 1/30, \mu_s^2 = 1/60$,
- Set 4: $c = 20, \lambda_1 = 0.2, \lambda_2 = 0.3, \mu_b = \mu_s^1 = 1/30, \mu_s^2 = 1/120$.

In Set 1 and Set 2, freight vehicles have a larger arrival rate than passenger vehicles; the reverse is the case in Set 3 and Set 4. We select identical parking rates for class-1 vehicles at the bay and street parking lots (i.e., $\mu_b = \mu_s^1$). In addition, the expected parking times of class-2 vehicles are set longer than the expected parking times of freight vehicles (i.e, $\mu_s^2 < \mu_b$). For time-dependent arrivals, we consider the following sets of parameters:

- Set 5: $c = 20, \overline{\lambda_1} = 0.4, \overline{\lambda_2} = 0.1, \alpha_1 = \alpha_2 = 0.5, \Psi_1 = 720, \Psi_2 = 1440, \mu_b = 1/30, \mu_s = 1/60$,
- Set 6: $c = 20, \overline{\lambda_1} = 0.4, \overline{\lambda_2} = 0.8, \alpha_1 = \alpha_2 = 0.5, \Psi_1 = 720, \Psi_2 = 1440, \mu_b = 1/30, \mu_s = 1/60$.

For these two sets of parameters, the cycle length of the passenger arrivals is equal to a full day of 24 hours and is double as the one of freight arrivals. The parameters α_1 and α_2 are set to 0.5, which captures a substantial time-dependent effect of +/-50% around the mean arrival rate. The mean arrival rates in Set 5 are chosen to represent a freight-intensive area while those of Set 6 represent a passenger-intensive area. In the approximation, we set the length of each interval L equal to 3 hours such that a full 24-hour cycle of arrivals for both passenger and freight vehicles is divided into 8 intervals. This value is observed to provide a good approximation in our experiments. The choice of L is in general complex to make and may require optimization. A small value for L is interesting as the mean arrival rate on each interval is close to the real value of the arrival rate at each point in time during the interval. However, with a small value of L, the stationary regime may not be reached on each interval and a non-negligible number of vehicles could be present on more than one interval which deteriorates the quality of the approximation.

Table 4 Maximum error

	Non-identical parking times				Time-dependent arrivals	
	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6
π_1^b	0.00%	0.00%	0.00%	0.00%	0.86%	0.86%
π_1	0.42%	0.56%	0.09%	0.23%	0.69%	0.76%
π_2	0.92%	1.14%	0.47%	0.82%	0.49%	0.09%
π	0.57%	0.72%	0.27%	0.54%	0.59%	0.27%
U_b	0.00%	0.00%	0.00%	0.00%	5.75%	5.75%
U_s	0.22%	0.21%	0.25%	0.15%	5.88%	0.30%

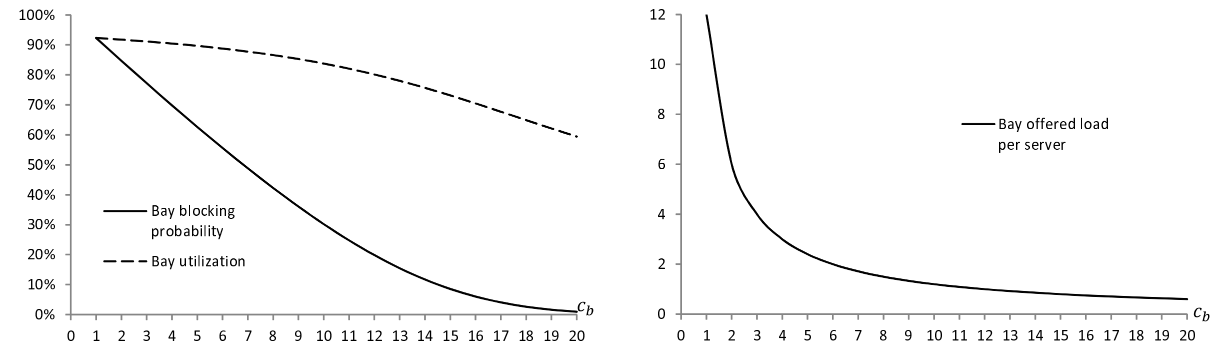
The results in Table 4 show the good quality of the approximations. For the approximation with non-identical parking times, the results with the approximation never exceed 1.2% difference with the simulation. Recall that at the bay parking stretch the approximation leads to the exact performance.

We also observe that the approximation in the case of time-dependent arrivals provides good results with a difference being below 1% for the blocking probabilities. Nevertheless, the bay and street utilizations show a larger gap between the approximation and the exact model. This observation is in line with the results in the literature explaining that the point-wise stationary approximation provides better results for extreme probabilities (e.g., probability of an empty system or blocking probability) than for metrics involving the distribution of the number of customers in the system (e.g., expected wait or expected number of customers) (Green and Kolesar 1991, Jennings et al. 1996). Yet, even for these metrics, the difference between the approximation and the simulation never exceeds 6%.

Another observation is that our approximations underestimate the simulations. This can be understood intuitively. For the approximation with non-identical parking times, the variability of the parking times is underestimated as we derive the performance measures under an exponential assumption whereas in the exact model the parking times have an hyperexponnetial distribution which has a higher variability than the exponential distribution. In the time-dependent approximation, we replace the real arrival rate by its mean on each interval. The resulting arrival process in the approximation then has a lower variability than the real one, leading to an underestimation of the performance measures.

6. Numerical evaluation

In this section, we perform a series of numerical experiments of our base model in order to obtain insights into the consequences of linking the bay parking and the street parking stretches into a single parking system. In particular, we characterize the impact of dedicating a certain number of on-street parking spaces to delivery bays on the fraction of lost freight (class-1) and passenger (class-2) vehicles (i.e., on their blocking probabilities), and on the utilization of the bay parking and the street parking stretches. Recall that we keep the total number of parking spaces constant such that if we allocate more parking spaces to the bay parking stretch, the number of parking spaces in the street parking stretch decreases by the same number. To provide broader general insights, we study both an area where more freight vehicles arrive than passenger vehicles (Section 6.1) and an area where the reverse is the case (Section 6.2). To facilitate the comparison, we select parameters such that the performance measures at the Bay parking are identical in these two examples (see Figure 2). In Section 6.3, we provide a practical illustration of the usage of our model using data from the Melbourne data set (Melbourne City Council 2014).



(a) Bay blocking probability (π_1^b) and bay utilization (U_b)   (b) Bay offered load per server (ρ_b)

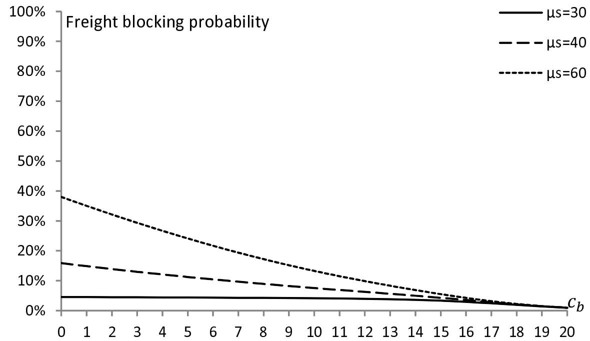
Figure 2   Bay performance measures for Examples 1 and 2 ($c = 20, \lambda_1 = 0.4, \mu_b = 1/30$**)**

6.1. Example 1: Freight-intensive area

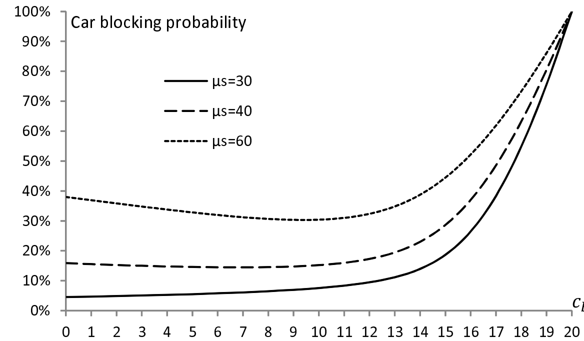
In this example, we have more freight vehicles arriving to the parking system than passenger vehicles (cars), as may be common in commercial downtown areas like a Central Business District. We set four times more freight vehicles arriving per time unit to the parking system than passenger vehicles. Freight vehicles park 30 minutes on average for unloading at the bay parking stretch. Parking times of passenger vehicles and freight vehicles arriving at the street parking stretch are assumed to be identical in distribution. For this setting, we choose the model parameters $c = 20, \lambda_1 = 0.4, \lambda_2 = 0.1, \mu_b = 1/30$. In this example, we vary the parking rate at the street parking stretch, μ_s, in order to gain insight into the impact of this parameter on performance measures of the parking system while also varying the number of parking spaces at the bay parking stretch, c_b.

The results are given in Figures 2 and 3. They illustrate Corollary 1. That is, the bay blocking probability (Figure 2(a)), the offered load per server to the bay parking stretch (Figure 2(b)), the fraction of lost freight vehicles (Figure 3(c)) and the system utilization (Figure 3(e)) decrease when the number of the bay parking spaces increases. We also observe that the bay utilization decreases in c_b (Figure 2(a)). The results involving freight vehicles are intuitive. Freight vehicles have access to more available parking spaces if more spaces are allocated to dedicated delivery bays. So, the freight blocking probability decreases in c_b (Figure 3(a)). It is also expected to have a decreasing system utilization in c_b as increasing c_b reduces the overall accessibility of the parking system.

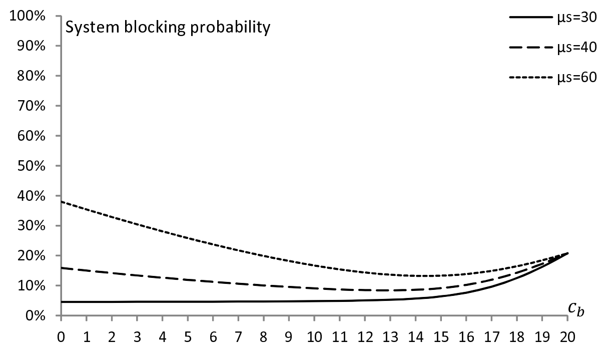
A non-expected result is that the fraction of lost cars (Figure 3(b)), the street utilization (Figure 3(d)) and the offered load per server at the street parking (Figure 3(f)) are not unimodal in c_b (i.e., strictly increasing or strictly decreasing). In particular, we observe that these different metrics have a minimum in c_b which differs from $c_b = 0$ (i.e., no dedicated bay parking space) or $c_b = c$ (i.e., inaccessible parking system to cars). This means that having an extra dedicated bay parking space does not necessarily deteriorate the service level for cars at the street parking stretch. It



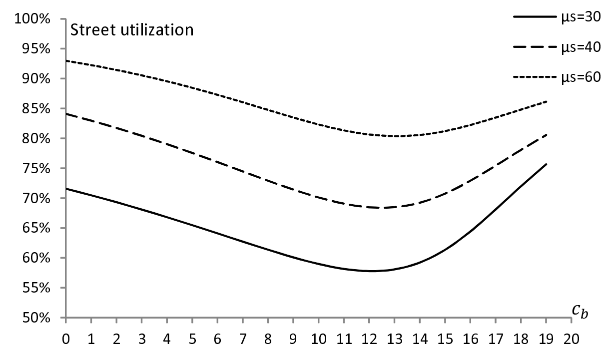
(a) Freight blocking probability (π_1)



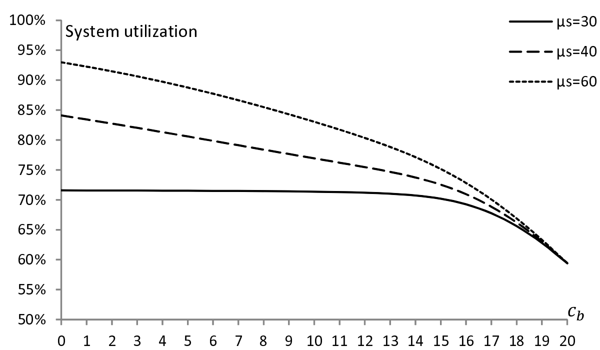
(b) Car blocking probability (π_2)



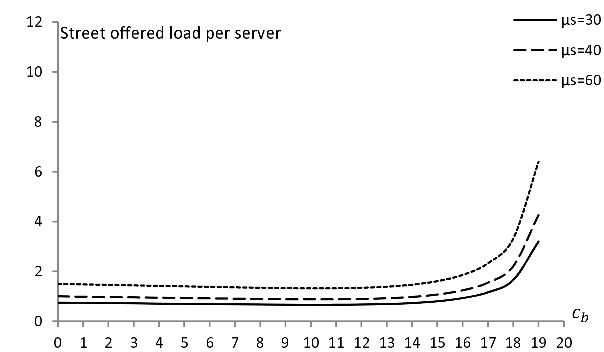
(c) System blocking probability (π)



(d) Street utilization (U_s)



(e) System utilization (U)



(f) Street offered load per server (ρ_s)

Figure 3 Street and system performance measures for Example 1 ($c = 20, \lambda_1 = 0.4, \lambda_2 = 0.1, \mu_b = 1/30, \mu_s = 1/30, 1/40, 1/60$**)**

may even reduce the car blocking probability. Two phenomena are in competition to explain this behavior. First, as c_b increases, c_s decreases due to $c = c_b + c_s$. Thus, cars have access to a smaller number of parking spaces which should *deteriorate* their service level. We call this phenomenon the *space-reduction effect*. Second, as c_b increases, freight vehicles have access to more dedicated parking spaces. Thus, the flow of freight vehicles at the street parking is reduced with c_b. This should *improve* the service level for cars as the competition for a parking space at the street parking is reduced when c_b increases. We call this phenomenon the *competition-reduction effect*. This existence of this second phenomenon is proven in Corollary 1 as the blocking probability at the bay parking is proven to decrease in c_b. We observe that the competition-reduction effect is more apparent when there is a high volume of freight vehicles compared to passenger vehicles as in the freight intensive area considered here ($\lambda_1 > \lambda_2$) and when cars spend a sufficiently long time at the parking (effect of μ_s).

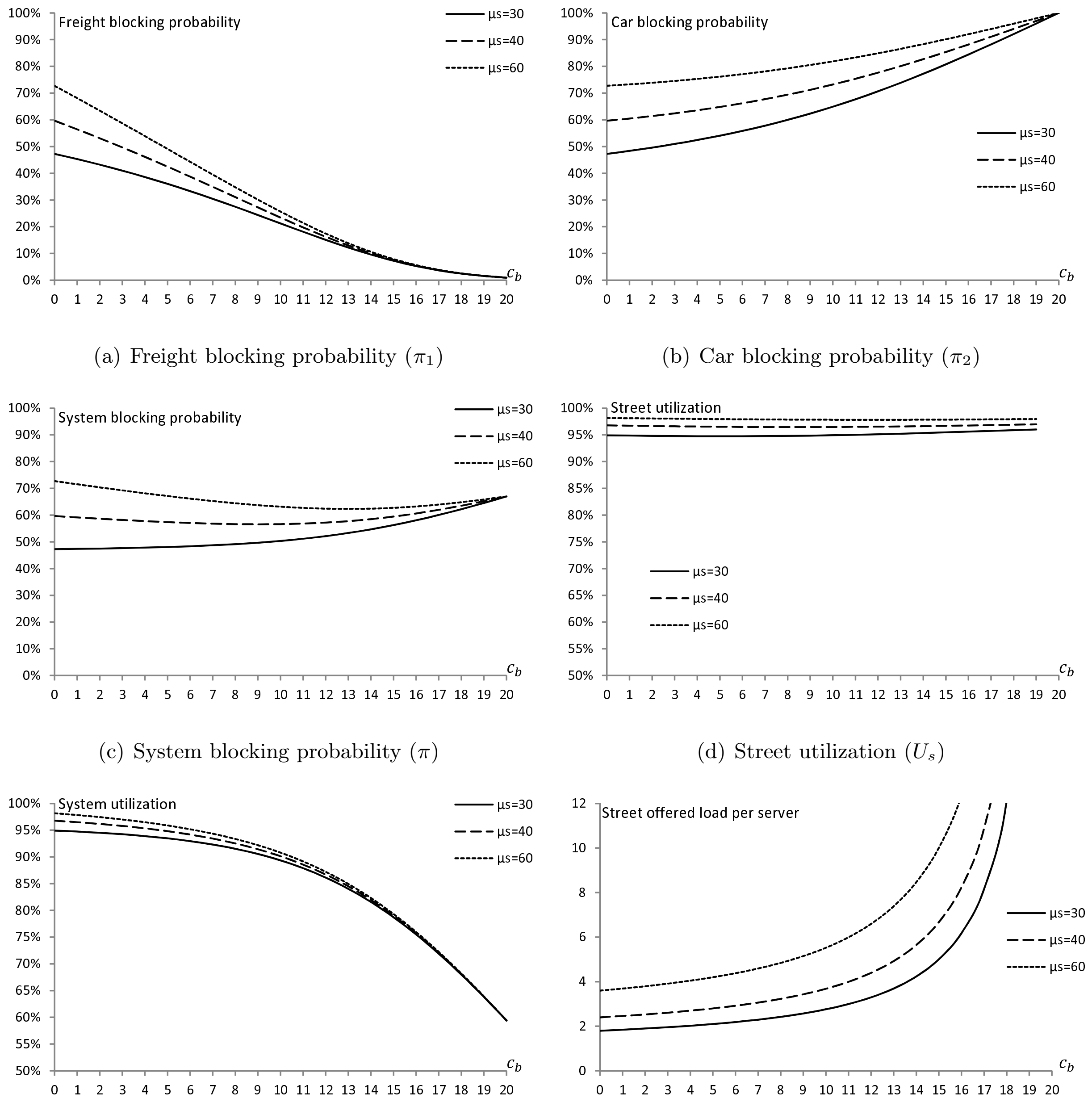
We note that the minimum of the street utilization and the offered load per server to the street parking stretch are not attained at the same size of the bay parking stretch c_b (and consequently also for the street parking stretch c_s). This is due to the fact that the behavior of the offered load per server does not change while changing the average parking time. As a consequence, it can be noticed from Table 5 that the minimum of the offered load per server is attained at $c_b = 10$, for all three scenarios: $\mu_s = 1/30$, $\mu_s = 1/40$ and $\mu_s = 1/60$. However, the street utilization is minimal when there are 12 parking spaces at the bay parking stretch ($c_b = 12$) for $\mu_s = 1/30$ and $\mu_s = 1/40$, and 13 parking spaces ($c_b = 13$) for $\mu_s = 1/60$. Due to the strong relationship between the freight vehicles lost to the bay parking stretch and the vehicles arriving to the street parking stretch, this system exhibits an interesting feature: it is possible to obtain a lower street utilization by having a higher offered load per server to the street parking stretch. In such case, it means that the competition-reduction effect is dominant for the street utilization while the space-reduction effect is dominant for the offered load per server to the street parking stretch.

Table 5 Street utilization and offered load per server to the street parking stretch ($c = 20, \lambda_1 = 0.4, \lambda_2 = 0.1, \mu_b = 1/30$)

		$\mu_s = 1/30$		$\mu_s = 1/40$		$\mu_s = 1/60$	
c_b	c_s	ρ_s	U_s	ρ_s	U_s	ρ_s	U_s
9	11	0.6659	0.6009	0.8879	0.7143	1.3318	0.8350
10	10	**0.6623**	0.5898	**0.8831**	0.7011	**1.3246**	0.8232
11	9	0.6637	0.5816	0.8849	0.6907	1.3274	0.8134
12	8	0.6729	**0.5779**	0.8971	**0.6848**	1.3457	0.8065
13	7	0.6941	0.5808	0.9255	0.6849	1.3882	**0.8038**
14	6	0.7344	0.5922	0.9792	0.6924	1.4688	0.8057

6.2. Example 2: Passenger-intensive area

In this example, the arrival rate of passenger vehicles is set at 0.8 (while it is 0.1 in Example 1), representing an area that is passenger-car intensive. In Figure 4, it can be noticed that, due to the higher arrival rate of passenger vehicles, the car blocking probability increases for all three scenarios ($\mu_s = 1/30$, $\mu_s = 1/40$ and $\mu_s = 1/60$) when the number of bay parking spaces increases (Figures



(a) Freight blocking probability (π_1)

(b) Car blocking probability (π_2)

(c) System blocking probability (π)

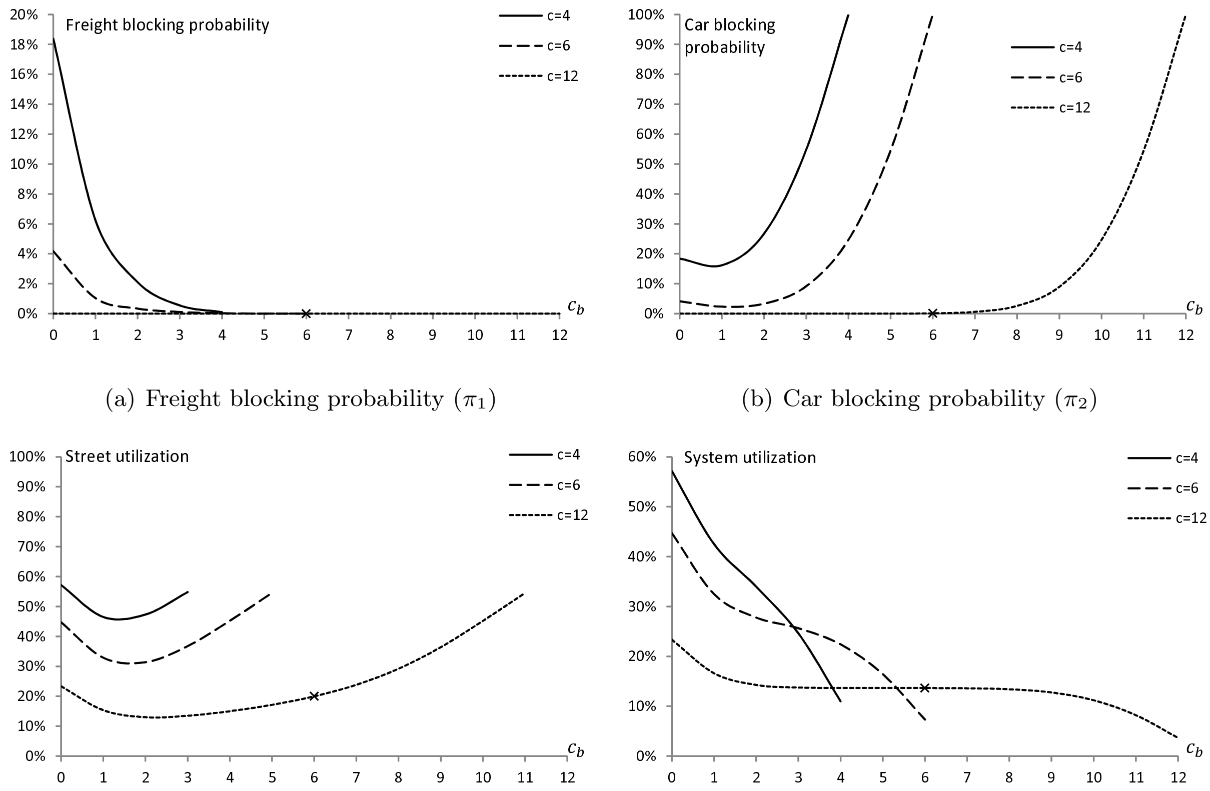
(d) Street utilization (U_s)

(e) System utilization (U)

(f) Street offered load per server (ρ_s)

Figure 4 Street and system performance measures for Example 2 ($c = 20, \lambda_1 = 0.4, \lambda_2 = 0.8, \mu_b = 1/30, \mu_s = 1/30, 1/40, 1/60$)

4(b), 4(f)). In this case, the space-reduction effect is dominant. Another interesting phenomenon is that the street utilization is insensitive to c_b (Figure 4(d)). This is a consequence of the fact that the arrival rate of cars is so high that the effect of decreasing the arrival rate of freight vehicles (that are blocked to the bay parking stretch) on the total arrival rate to the street parking stretch is negligible (i.e., negligible competition-reduction effect). As a result, it becomes harder for freight vehicles to park at the street parking stretch.



(a) Freight blocking probability (π_1)

(b) Car blocking probability (π_2)

(c) Street utilization (U_s)

(d) System utilization (U)

Figure 5 Street and system performance measures for Example 3 ($c = 4, 6, 12, \lambda_1 = 0.04, \lambda_2 = 0.03, \mu_b = 1/11, \mu_s = 1/40$**)**

6.3. Example 3: Parameters based on empirical data

In this example, we demonstrate our methodology with parameters based on the empirical data that we used above to support our assumption that the arrival processes of both freight and passenger vehicles follow Poisson processes with different rate parameters, and their parking times are exponentially distributed. We use these 2014 data from the city of Melbourne, Australia (Melbourne City Council 2014) that have been collected using sensors on parking bays in the downtown area. In addition, the dataset includes specific information whether a parking bay is a delivery bay (for freight) or a general parking area (for general on-street parking). We use all transactional data for one particular parking system consisting of a stretch of 6 delivery bays and 6 general on-street parking places, and use the sample means as estimators for the model parameters.

The estimated model parameters are $c = 12, c_b = 6, \lambda_1 = 0.04, \lambda_2 = 0.03, \mu_b = 1/11, \mu_s = 1/40$. Figure 5 summarizes the main performance metrics for this Melbourne example as functions of c_b. The position which corresponds to $c = 12$ and $c_b = 6$ is given by a cross on the curve corresponding to $c = 12$. We also add the performance measures for $c = 4$ and $c = 6$ to illustrate the impact of reducing the number of parking spaces. In this example, the system utilization of the parking system in Melbourne is considerably lower than the settings we studied in the previous two examples (Figure 5(d)). Consequently, the vehicle blocking probabilities are very low (Figures 5(a) and 5(b)). Actually, from an urban planning perspective, a very similar performance could be obtained with only half size of the parking system (6 spaces), of which 2 are allocated to delivery bays. Even with only 4 spaces available, the car blocking probability could be limited to under 0.2 and the freight vehicle blocking probability under 0.06 if the number of delivery bays is set to 1 (out of 4 available spaces). Interestingly, this specific numerical study shows how our model and associated insights could be used to assess the overall sizes of parking systems, in addition to the earlier illustrated decision support for the allocation decision.

7. Discussion and concluding remarks

In dense urban environments, freight delivery is highly fragmented. Many small stores can be served on a route in emerging economies, while many homes can be served in a route in the most developed cities of this world. Delivery vehicles to mom-and-pop operated nanostores in cities such as Bogota or Quito have been reported to visit more than 100 such stores on a route, and courier companies in cities such as London or Paris also make more than 150 stops on a route. As a result, vehicles tend to be parked for a significant share of the route. In cities, this is visible on a day-to-day basis by double-parked or otherwise illegally parked freight vehicles. Due to delivery times being highly stochastic, both the parking times and the interarrival times of vehicles are also stochastic. It hence makes sense to develop models that enable us to evaluate the performance of the allocation of public space to bays that are dedicated for conducting deliveries.

In this paper, we analyze such a parking system as a queueing model with a stretch that is reserved exclusively for freight deliveries and a general purpose on-street parking stretch. They are modeled as a set of parallel servers. Upon arrival, delivery vehicles first attempt to park in a free dedicated delivery bay; if not available, they park on the street. If also no street parking is available, they leave our system not to return; in practice such blocking will lead to illegal parking as deliveries always take place, and hence the freight vehicle system blocking probability can be seen as the probability to park illegally. We believe our model is the first to study the role of delivery bays in an urban logistics setting in a stochastic manner. The framework we provide can serve as a basis for further work in this area, as it can be extended to a queueing network to represent more extensive relations between multiple separate bay parking stretches and street parking stretches in an urban setting. Our modeling approach hence provides a basis for much-needed analysis of the usage of scarce public space for delivery in dense cities. In particular the fact that freight vehicles make use of the general on-street parking if delivery bays are not available, leads to intricate behavior of the entire parking system.

The analysis of our model creates parsimonious insights into the behavior of a delivery bay parking stretch as part of a limited stretch of curbside. We have provided explicit expressions for the relevant performance measures, and formally proven a number of monotonicity results. Our numerical results show that increasing the share of delivery bays decreases the share of delivery vehicles that is lost to our system in line with decreased utilization of the bays. However, the effect on the fraction of cars that is lost does not necessarily display such monotonic behavior. Especially in areas that are freight-intensive in terms of curbside parking, also passenger vehicles might be better off if more spaces are allocated for freight delivery only. Further, we illustrate for one particular real-life parking system that a very similar performance in terms of blocked vehicles could be obtained with only half the number of parking spaces, freeing up public space that can be appropriate for other purposes.

Our modeling approach can serve as a basis for more extensive models. For instance, it can be enriched by studying interactions between different parking stretches, as for instance passenger vehicles may be inclined to seek parking further away if no space is available. Also, it can be interesting to study temporal allocation of delivery bays in case of time-dependent arrivals of freight vehicles.

Acknowledgments

The authors thank Marko Jakšič for his valuable input on earlier drafts of this paper and Juan Muriel Villegas for sharing the Melbourne parking data. We also would like to thank the Associate Editor and Reviewer team for their detailed and constructive reviews that have helped us improving the paper.

References

Bassamboo A, Harrison JM, Zeevi A, 2006 *Design and control of a large call center: Asymptotic analysis of an LP-based method. Operations Research* 54(3):419–435.

Bhulai S, Brooms AC, Spieksma FM, 2014 *On structural properties of the value function for an unbounded jump markov process with an application to a processor sharing retrial queue. Queueing Systems* 76(4):425–446.

Caliskan M, Barthels A, Scheuermann B, Mauve M, 2007 *Predicting parking lot occupancy in vehicular ad hoc networks. 2007 IEEE 65th Vehicular Technology Conference-VTC2007-Spring*, 277–281 (IEEE).

Cassady CR, Kobza JE, 1998 *A probabilistic approach to evaluate strategies for selecting a parking space. Transportation Science* 32(1):30–42.

Crainic TG, Ricciardi N, Storchi G, 2009 *Models for evaluating and planning city logistics systems. Transportation Science* 43(4):432–454.

Daley D, 1976 *Queueing output processes. Advances in Applied Probability* 8(2):395–415.

Dalla Chiara G, Cheah L, 2017 *Data stories from urban loading bays. European Transport Research Review* 9(4):50.

Dalla Chiara G, Goodchild A, 2020 *Do commercial vehicles cruise for parking? Empirical evidence from Seattle. Transport Policy* 97:26–36.

Dowling CP, Ratliff LJ, Zhang B, 2019 *Modeling curbside parking as a network of finite capacity queues. IEEE Transactions on Intelligent Transportation Systems* 21(3):1011–1022.

Eick SG, Massey WA, Whitt W, 1993 $M_t/G/\infty$ *queues with sinusoidal arrival rates. Management Science* 39(2):241–252.

Fransoo JC, Cedillo-Campos MG, Gamez-Perez KM, 2020 *Estimating the benefits of dedicated unloading bays by field experimentation.* Working paper, Tilburg University.

Gans N, Zhou Y, 2007 *Call-routing schemes for call-center outsourcing. Manufacturing & Service Operations Management* 9(1):33–50.

Gao J, Ozbay K, 2016 *Modeling double parking impacts on urban street. Transportation Research Board 95th Annual Meeting*, volume 12.

Goodchild A, Ivanov B, 2017 *The final 50 feet of the urban goods delivery system.* Technical report, University of Washington.

Green L, Kolesar P, 1991 *The pointwise stationary approximation for queues with nonstationary arrivals. Management Science* 37(1):84–97.

Gurvich I, Perry O, 2012 *Overflow networks: Approximations and implications to call center outsourcing. Operations Research* 60(4):996–1009.

Hauer E, Templeton JGC, 1972 *Queuing in lanes. Transportation Science* 6(3):247–259.

Jaller M, Holguín-Veras J, Hodge SD, 2013 *Parking in the city: Challenges for freight traffic. Transportation Research Record* 2379(1):46–56.

Jennings O, Mandelbaum A, Massey W, Whitt W, 1996 *Server staffing to meet time-varying demand. Management Science* 42(10):1383–1394.

Kladeftiras M, Antoniou C, 2013 *Simulation-based assessment of double-parking impacts on traffic and environmental conditions. Transportation Research Record* 2390(1):121–130.

Kolmogorov A, 1933 *Sulla determinazione empirica di una legge di distribuzione. Giornale dell'Istituto Italiano degli Attuari* 4:83–91.

Koole G, Nielsen B, Nielsen T, 2015 *Optimization of overflow policies in call centers. Probability in the Engineering and Informational Sciences* 29(3):461–471.

Ku C, Jordan S, 2003 *Near optimal admission control for multiserver loss queues in series. European Journal of Operational Research* 144(1):166–178.

Kuczura A, 1973 *Loss systems with mixed renewal and poisson inputs. Operations Research* 21(3):787–795.

Larson RC, Sasanuma K, 2007 *Congestion pricing: A parking queue model.* Working paper, Massachusetts Institute of Technology. Engineering Systems Division.

Legros B, 2021 *Dimensioning a queue with state-dependent arrival rates. Computers & Operations Research* 128:105–179.

Legros B, Jouini O, 2019 *On the scheduling of operations in a chat contact center. European Journal of Operational Research* 274(1):303–316.

Little JD, 1961 *A proof for the queuing formula: L= λ W. Operations Research* 9(3):383–387.

Maglaras C, Van Mieghem J, 2005 *Queueing systems with leadtime constraints: A fluid-model approach for admission and sequencing control. European journal of Operational Research* 167(1):179–207.

Malik L, Sánchez-Díaz I, Tiwari G, Woxenius J, 2017 *Urban freight-parking practices: The cases of Gothenburg (Sweden) and Delhi (India). Research in Transportation Business & Management* 24:37–48.

Melbourne City Council, 2014 *Parking bay arrivals and departures 2014.* Retrieved from `https://data.melbourne.vic.gov.au/Transport/Parking-bay-arrivals-and-departures-2014/mq3i-cbxd`.

Niyirora J, Zhuang J, 2017 *Fluid approximations and control of queues in emergency departments. European Journal of Operational Research* 261(3):1110–1124.

Nourinejad M, Wenneman A, Habib KN, Roorda MJ, 2014 *Truck parking in urban areas: Application of choice modelling within traffic microsimulation. Transportation Research Part A: Policy and Practice* 64:54–64.

Puterman ML, 2014 *Markov decision processes: Discrete stochastic dynamic programming* (John Wiley & Sons).

Ross SM, 2014 *Introduction to probability models* (Academic press).

Savelsbergh M, Van Woensel T, 2016 *50th anniversary invited article—city logistics: Challenges and opportunities. Transportation Science* 50(2):579–590.

Shoup DC, 2006 *Cruising for parking. Transport Policy* 13(6):479–486.

Smirnov N, 1948 *Table for estimating the goodness of fit of empirical distributions. Annals of Mathematical Statistics* 19(2):279–281.

Takács L, 1959 *On the limiting distribution of the number of coincidences concerning telephone traffic. Annals of Mathematical Statistics* 30(1):134–142.

Taniguchi E, Thompson RG, 2018 *City Logistics 1: New Opportunities and Challenges* (John Wiley & Sons).

Vieira JGV, Fransoo JC, 2015 *How logistics performance of freight operators is affected by urban freight distribution issues. Transport Policy* 44:37–47.

Ward A, Kumar S, 2008 *Asymptotically optimal admission control of a queue with impatient customers. Mathematics of Operations Research* 33(1):167–202.

Wigan MR, Broughton J, 1980 *Descriptive models of parking to complement transport planning studies. Transportation Research Part A: General* 14(3):159–171.

Xiao J, Lou Y, Frisby J, 2018 *How likely am I to find parking? A practical model-based framework for predicting parking availability. Transportation Research Part B: Methodological* 112:19–39.

Xu K, 2015 *Necessity of future information in admission control. Operations Research* 63(5):1213–1226.

Appendix A: Proof of Theorem 1

Since $V_0(c_b,x,y)=0$, $V_0\in\mathcal{C}$ and $V_0\in\mathcal{C}'$. Let us assume that $V_k\in\mathcal{C}'$. We want to show that $V_{k+1}\in\mathcal{C}'$.

Relation (10). For $0\le x\le c_b\le c$, and $0\le y\le c-(c_b+1)$, we have

$$\begin{aligned}
V_{k+1}(c_b,x,y)-V_{k+1}(c_b+1,x,y) &= f(c_b,x,y)-f(c_b+1,x,y)\\
&+\lambda_1\,(\mathbb{1}_{x<c_b}\,[V_k(c_b,x+1,y)-V_k(c_b+1,x+1,y)]\\
&+\mathbb{1}_{x=c_b,y<c-c_b}\,[V_k(c_b,x,y+1)-V_k(c_b+1,x+1,y)])\\
&+\lambda_2\left(\mathbb{1}_{y<c-(c_b+1)}\,[V_k(c_b,x,y+1)-V_k(c_b+1,x,y+1)]\right.\\
&\left.+\mathbb{1}_{y=c-(c_b+1)}\,[V_k(c_b,x,y+1)-V_k(c_b+1,x,y)]\right)\\
&+x\mu_b\,[V_k(c_b,x-1,y)-V_k(c_b+1,x-1,y)]+y\mu_s\,[V_k(c_b,x,y-1)-V_k(c_b+1,x,y-1)]\\
&+(1-\lambda_1-\lambda_2-x\mu_b-y\mu_s)\,[V_k(c_b,x,y)-V_k(c_b+1,x,y)]\,.
\end{aligned}$$

Each line on the right hand side of this equation is positive. Lines 1, 2, 4, 6, and 7 are positive since f and V_k satisfy (10). Line 3 is positive since V_k satisfies (13). Finally, Line 5 is positive since V_k satisfies (14). This proves that V_{k+1} satisfies (10).

Relation (11). For $0\le x\le c_b\le c$, and $0\le y\le c-(c_b+1)$, we have

$$\begin{aligned}
V_{k+1}(c_b,x,y+1)-V_{k+1}(c_b,x,y) &= f(c_b,x,y+1)-f(c_b,x,y)\\
&+\lambda_1\,(\mathbb{1}_{x<c_b}\,[V_k(c_b,x+1,y+1)-V_k(c_b,x+1,y)]\\
&+\mathbb{1}_{x=c_b,y<c-(c_b+1)}\,[V_k(c_b,x,y+2)-V_k(c_b,x,y+1)]\\
&+\mathbb{1}_{x=c_b,y=c-(c_b+1)}\,[V_k(c_b,x,y+1)-V_k(c_b,x,y+1)])\\
&+\lambda_2\left(\mathbb{1}_{y<c-(c_b+1)}\,[V_k(c_b,x,y+2)-V_k(c_b,x,y+1)]+\mathbb{1}_{y=c-(c_b+1)}\,[V_k(c_b,x,y+1)-V_k(c_b,x,y+1)]\right)\\
&+x\mu_b\,[V_k(c_b,x-1,y+1)-V_k(c_b,x-1,y)]+y\mu_s\,[V_k(c_b,x,y)-V_k(c_b,x,y-1)]+\mu_sV_k(c_b,x,y)\\
&+(1-\lambda_1-\lambda_2-x\mu_b-(y+1)\mu_s)\,[V_k(c_b,x,y+1)-V_k(c_b,x,y)]-\mu_sV_k(c_b,x,y).
\end{aligned}$$

Each line on the right hand side of this equality is positive as f and V_k satisfy (11). Note that the third line is zero and that the terms proportional with μ_s at the two last lines sum up to zero. This proves that V_{k+1} satisfies (11).

Relation (13). For $0\le x\le c_b\le c$, and $0\le y\le c-(c_b+1)$, we have

$$\begin{aligned}
V_{k+1}(c_b,x,y+1)-V_{k+1}(c_b+1,x+1,y) &= f(c_b,x,y+1)-f(c_b+1,x+1,y)\\
&+\lambda_1\,(\mathbb{1}_{x<c_b}\,[V_k(c_b,x+1,y+1)-V_k(c_b+1,x+2,y)]\\
&+\mathbb{1}_{x=c_b,y<c-(c_b+1)}\,[V_k(c_b,x,y+2)-V_k(c_b+1,x+1,y+1)]\\
&+\mathbb{1}_{x=c_b,y=c-(c_b+1)}\,[V_k(c_b,x,y+1)-V_k(c_b+1,x+1,y)])
\end{aligned}$$

$$
\begin{aligned}
&+\lambda_2\left(\mathbb{1}_{y<c-(c_b+1)}\left[V_k(c_b,x,y+2)-V_k(c_b+1,x+1,y+1)\right]\right.\\
&\left.+\mathbb{1}_{y=c-(c_b+1)}\left[V_k(c_b,x,y+1)-V_k(c_b+1,x+1,y)\right]\right)\\
&+x\mu_b\left[V_k(c_b,x-1,y+1)-V_k(c_b+1,x,y)\right]-\mu_b V_k(c_b+1,x,y)\\
&+y\mu_s\left[V_k(c_b,x,y)-V_k(c_b+1,x+1,y-1)\right]+\mu_s V_k(c_b,x,y)\\
&+(1-\lambda_1-\lambda_2-(x+1)\mu_b-(y+1)\mu_s)\left[V_k(c_b,x,y+1)-V_k(c_b+1,x+1,y)\right]\\
&+\mu_b V_k(c_b,x,y+1)-\mu_s V_k(c_b+1,x+1,y)\\
&\geq\mu_b\left(V_k(c_b,x,y+1)-V_k(c_b+1,x,y)\right)+\mu_s(V_k(c_b,x,y)-V_k(c_b+1,x+1,y))\\
&=(\mu_b-\mu_s)\left(V_k(c_b,x,y+1)-V_k(c_b+1,x,y)\right)\\
&+\mu_s(\left[V_k(c_b,x,y)-V_k(c_b+1,x,y)\right]+\left[V_k(c_b,x,y+1)-V_k(c_b+1,x+1,y)\right])
\end{aligned}
$$

All terms on the right hand side of the equality are positive since they all satisfy (13). The remaining terms proportional with μ_s and μ_b are kept after the inequality sign. Next, by rewriting these elements into a first part proportional with $\mu_b-\mu_s$ and another part proportional with μ_s, we prove that V_{k+1} satisfies (13). The term proportional with $\mu_b-\mu_s$ is positive as V_k satisfies (14). The term proportional with μ_s is also positive since V_k satisfies (10) and (13).

This proves that if $V_k\in\mathcal{C}'$ and $\mu_b\geq\mu_s$, then $V_{k+1}\in\mathcal{C}'$. The condition $\mu_b\geq\mu_s$ can be seen as a restriction of our result. Alternatively, we can show (12) independently from any conditions on the system parameters. However, this relation is rarely satisfied by the cost functions.

Relation (12). For $0\leq x\leq c_b\leq c$, and $0\leq y\leq c-(c_b+1)$, we have

$$
\begin{aligned}
V_{k+1}(c_b,x,y)-V_{k+1}(c_b+1,x+1,y)&=f(c_b,x,y)-f(c_b+1,x+1,y)\\
&+\lambda_1\left(\mathbb{1}_{x<c_b}\left[V_k(c_b,x+1,y)-V_k(c_b+1,x+2,y)\right]\right.\\
&+\mathbb{1}_{x=c_b,y<c-(c_b+1)}\left[V_k(c_b,x,y+1)-V_k(c_b+1,x+1,y+1)\right]\\
&\left.+\mathbb{1}_{x=c_b,y=c-(c_b+1)}\left[V_k(c_b,x,y+1)-V_k(c_b+1,x+1,y)\right]\right)\\
&+\lambda_2\left(\mathbb{1}_{y<c-(c_b+1)}\left[V_k(c_b,x,y+1)-V_k(c_b+1,x+1,y+1)\right]\right.\\
&\left.+\mathbb{1}_{y=c-(c_b+1)}\left[V_k(c_b,x,y+1)-V_k(c_b+1,x+1,y)\right]\right)\\
&+x\mu_b\left[V_k(c_b,x-1,y)-V_k(c_b+1,x,y)\right]\\
&+y\mu_s\left[V_k(c_b,x,y-1)-V_k(c_b+1,x+1,y-1)\right]\\
&+(1-\lambda_1-\lambda_2-(x+1)\mu_b-y\mu_s)\left[V_k(c_b,x,y)-V_k(c_b+1,x+1,y)\right]\\
&+\mu_b\left(V_k(c_b,x,y)-V_k(c_b+1,x,y)\right).
\end{aligned}
$$

Again, the right hand side of this Equation is positive. Lines 1, 2, 3, 5, 7, 8, and 9 are positive since V_k satisfies (12). Lines 4 and 6 are positive since V_k satisfies (13). Line 10 is positive since V_k satisfies (10). This proves that V_{k+1} satisfies (12) and finishes the proof of the theorem.